AD_____

Award Number: W81XWH-10-1-0519

TITLE: Massively Parallel Rogue Cell Detection Using Serial Time-Encoded Amplified Microscopy of Inertially Ordered Cells in High-Throughput Flow

PRINCIPAL INVESTIGATOR: Dr. Dino Di Carlo

CONTRACTING ORGANIZATION: University of California, Los Angeles
Los Angeles, CA 90095

REPORT DATE: August 2011

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-08-2011	Annual	1 AUG 2010 - 31 JUL 2011

4. TITLE AND SUBTITLE
Massively Parallel Rogue Cell Detection Using Serial Time-Encoded Amplified Microscopy of Inertially Ordered Cells in High-Throughput Flow

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-10-1-0519

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Dino Di Carlo

E-Mail: dicarlo@seas.ucla.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California, Los Angeles
Los Angeles, CA 90095

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We are developing a high-throughput imaging cytometry system for high accuracy identification of rare circulating breast cancer cells in blood. Key accomplishments to date include: 1. Construction of the optical system including the STEAM camera and its integration with the microfluidic device as well as the optical image amplifier and serializer for high-speed blur-free image acquisition, 2. Development of the microfluidic device made of rigid thermoset polyester, 3. Development of the real-time optoelectronic image processor that performs high-throughput image-based screening of a large number of cells and identification of rare cells, 4. Development of the algorithms implemented on the FPGA and CPU, and 5. Integration of the hardware and software into a combined system. These accomplishments have led to a successful demonstration of the STEAM flow analyzer for real-time detection of rare MCF7 cells in lysed blood with sensitivity of 75% and a false positive rate of 1 in 1,000,000 white blood cells.

15. SUBJECT TERMS
Microfluidics, circulating tumor cells, rare cell detection, high-speed imaging, flow cytometry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	13	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

We aim to develop an instrument for high-throughput identification of rare circulating breast cancer cells to enable early detection and analysis of treatment effectiveness. To address this challenge, we have developed an automated flow-through single-cell optical microscopy system that can evaluate, diagnose, and screen a large population of cells in a short time. This method builds on an unique integration of (i) an ultrafast optical imaging modality known as serial time-encoded amplified microscopy (STEAM) [1] for blur-free imaging of cells in high-speed flow, (ii) inertial microfluidic technology for sheath-free focusing and ordering of cells with inertial forces [2-4], and (iii) hybrid optoelectronic image processing circuitry for real-time image processing. The integrated system transforms microfluidic flow into a series of "E-slides" – an electronic version of glass slides – on which cells of interest are digitally analyzed. This property enables fully automated real-time image-recording and classification of a large number of cells through their morphological and biochemical features. As our first proof-of-principle demonstration, we have shown non-stop real-time image-based identification and screening of rare MCF7 breast cancer cells in blood with an unprecedented throughput of 100,000 cells/s and false positive rate of 1 in a million.

Body

Our system, which we refer to as the STEAM flow analyzer (Figure 1, Supplementary Table 1), consists of three subsystems: (i) the microfluidic device (Figure 2, Supplementary Table 2), (ii) the STEAM camera (Figure 1, Supplementary Table 1), and (iii) the real-time optoelectronic image processor (Figure 3, Supplementary Table 3). The STEAM flow analyzer operates in three steps. First, cells are controlled to flow at a uniform velocity and focused and ordered by inertial lift forces in the microfluidic channel. Second, the STEAM camera takes images of the fast-flowing cells. Finally, the real-time optoelectronic image processor optoelectronically processes the images and performs automated cell screening in real time.

The STEAM flow analyzer employs inertial microfluidic technology (Figure 2). This method provides uniform positions and velocities for cells with inertial lift forces (important for generation of clear E-slides) while eliminating the need for sheath fluid which is traditionally required for hydrodynamic focusing in conventional flow cytometers. To ensure stability in real-time image processing, the microfluidic device was fabricated by replica molding in thermoset polyester due to its high stiffness and ability to sustain high pressures (Appendices, Supplementary Figure 1, Supplementary Table 2).

The STEAM camera operates in a shadowgraph imaging configuration. First, the image of a cell is encoded into the spectrum of broadband optical pulses. Operating the STEAM camera in a line scan mode and double-pass shadowgraph imaging configuration, the real-time optoelectronic image processor (the details of which are described below) constructs the 2D shadowgraph image of a cell from a series of 1D line scans (Figure 1). For the line scans, a pair of diffraction gratings spatially disperses and hence converts broadband optical pulses into rainbow pulses (27 ps pulse width) which are incident via the objective lens onto the flowing cell in the microfluidic channel. The pulses are repeated at a repetition rate of 36.7 MHz, which corresponds to the line scan rate. The back-reflected pulses from the substrate of the microfluidic device are directed via the same optics and an optical circulator toward the real-time optoelectronic image processor in which 2D E-slides are constructed from the 1D frames and subject to screening.

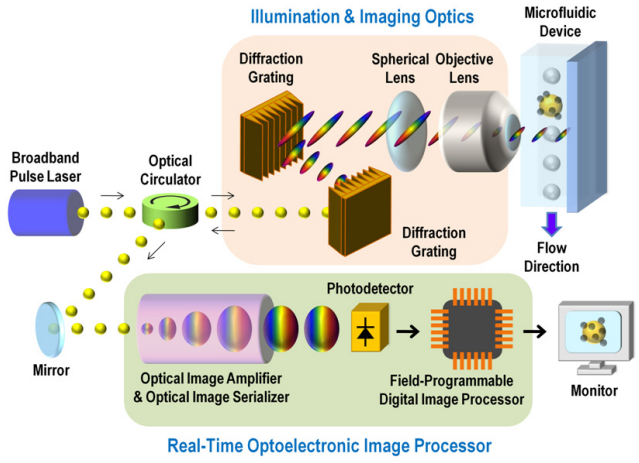


Figure 1: Schematic of the STEAM flow analyzer that highlights an optical layout of the STEAM camera and real-time optoelectronic image processor. The STEAM camera takes blur-free images of fast-flowing cells in the microfluidic device. The acquired images are optoelectronically processed and screened in the real-time image processor.

The architecture and function of real-time optoelectronic image processor are multi-fold. As shown in Figure 1, it consists of an amplified dispersive Fourier transformer (i.e., optical image amplifier and serializer), a high-speed photodetector, and a field-

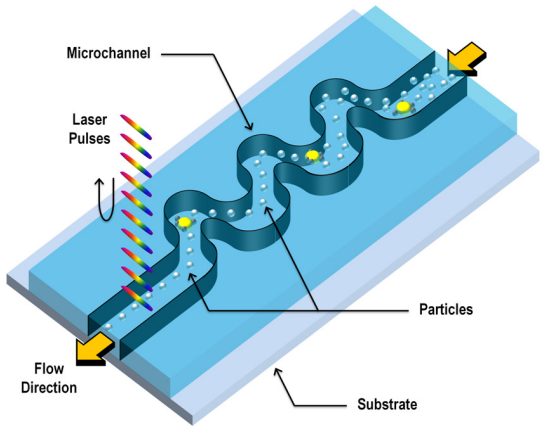


Figure 2: Microfluidic device in which cells are controlled to flow at a uniform velocity and focused and ordered by inertial lift forces in the microfluidic channel.

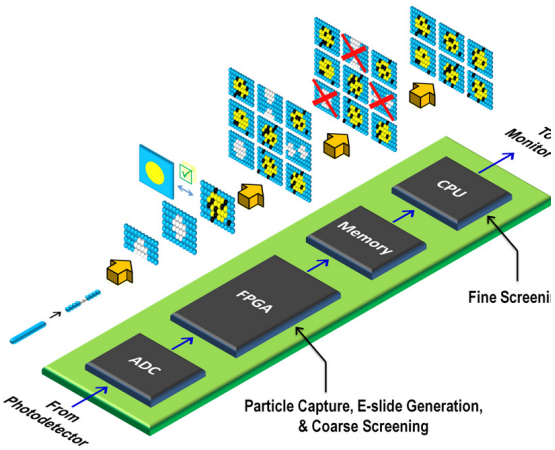


Figure 3: Field-programmable digital image processor that consists of a high-speed analog-to-digital converter (ADC), field-programmable gate array (FPGA), memory circuit, and central processing unit (CPU). It captures cells in the field of view, generates E-slides for the cells, and performs fully automated screening in real time.

programmable digital image processor. First, the amplified dispersive Fourier transformer serializes the image-encoded pulse (i.e., the 1D image frame) into a time-domain data stream and simultaneously amplifies it in the optical domain (i.e., optical image serialization and amplification). The optical image amplification (~30 dB) before photon-to-electron conversion is critical as it enables high-throughput microscopy in the ultrashort exposure time (27 ps) without the need for a high-intensity illuminator or sacrificing sensitivity. In addition, the ultrafast shutter speed freezes any motion of cells in high-speed flow, allowing for acquisition of blur-free images. The amplified 1D data stream is detected by the high-speed single-pixel photodetector, eliminating the need for a detector array and hence readout time limitations. Second, the field-programmable digital image processor (Figure 3, Appendices) employs a programmable multi-stage decision-making architecture that can be customized, depending on the application. It consists of (i) a high-speed analog-to-digital converter (ADC) for image digitization, (ii) a fully integrated designer-programmable circuit employing a field-programmable gate array (FPGA) for cell capture, E-slide generation, and coarse cell classification, (iii) a memory circuit for storing down-selected E-slides, and (iv) a central processing unit (CPU) for fine cell classification. Specifically, the function of the FPGA is to identify the presence of cells (ignoring the gaps between the cells), generate E-slides for the cells, down-select the E-slides (i.e., the cells) through their morphological and biochemical markers, and store the selected E-slides in the memory while dumping the unselected E-slides. To ensure reliable cell classification on the field-programmable digital image processor, our in-house cell classification method (the so-called custom supervised learning method or support vector machine) is trained to build a model that predicts whether a new target (i.e., cell) falls into any of the

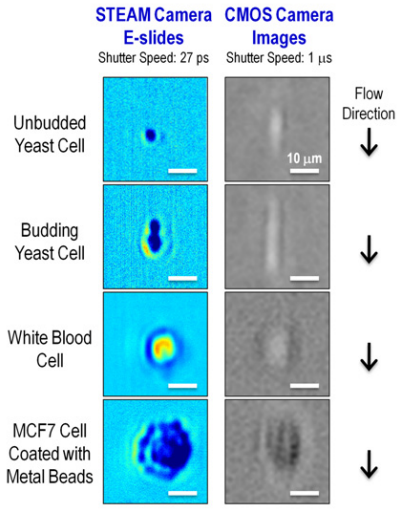


Figure 4: Performance of the STEAM camera and comparison with a state-of-the-art CMOS camera. E-slides of flowing cells of various species in the microfluidic device were generated by the STEAM flow analyzer with the built-in STEAM camera (27 ps shutter speed, 128 x 512 pixels, 25 dB optical image gain). The E-slides are compared with images of the same particles captured by a state-of-the-art CMOS camera (1 µs shutter speed, 32 x 32 pixels) under the same flow conditions. To operate at these high speeds, the CMOS camera used partial readout, limiting the number of pixels to 32 x 32. Here the cells were controlled to flow at a uniform speed of 4 m/s, which corresponds to a throughput of 100,000 cells/s based on the volumetric flow rate. The high-speed motion of the particles was frozen by the ultrafast shutter speed of the STEAM camera yet without sacrificing sensitivity due to the optical image amplification whereas the reduced sensitivity and motion blurs caused by the CMOS camera's much lower shutter speed and lack of optical image amplification are evident.

pre-defined cell categories, and then used to identify the target cells in real time (Appendices).

The STEAM flow analyzer's blur-free image acquisition enables differentiation of cells from a heterogeneous population. Figure 4 shows E-slides of fast-flowing cells of various species in the microfluidic channel captured by the STEAM flow analyzer. Here the cells were controlled to flow at a uniform speed of 4 m/s, which corresponds to a throughput of 100,000 cells/s – a very fast flow, but motion blur is absent in the images due to the ultrafast shutter speed of the STEAM camera (27 ps). For comparison, Figure 4 also shows images of the same types of cells flowing under the same conditions captured by a state-of-the-art CMOS camera. These images show that the CMOS camera's lower shutter speed (1 μs) and lack of optical image amplification significantly reduced the sensitivity and caused motion blur in the images, rendering the camera unable to classify cells reliably. For further comparison, stationary cells of the same types on a glass slide were obtained under a conventional microscope with a CCD camera with a much longer exposure time (shutter speed) of 17 ms (Supplementary Figure 2). Despite the fact that the STEAM camera is many orders of magnitude faster than the CCD camera, the two cameras share similar image quality (i.e., sensitivity and resolution).

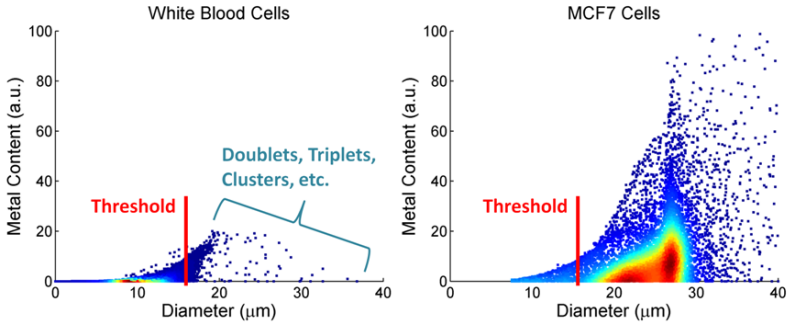


Figure 5: Scatter plots of white blood cells and MCF7 cells based on the cell size (i.e., area) and the presence of the surface antigens (i.e., metal beads). The total number of events is ~10,000 for both cell types. This statistical analysis is used to build a model and train the supervised learning method and hence the algorithms to run the field-programmable digital image processor. The threshold for the size-based selection performed on the FPGA is set such that smaller MCF7 cells are also selected at the expense of detecting larger white blood cells. The rare white blood cell events that overlap with the distribution of MCF7 cells are doublets or clusters and can hence be rejected by imaging (which is not possible with single-point detection methods).

To show the utility of the STEAM flow analyzer, we used it to demonstrate detection of rare breast cancer cells in blood. Such rare cells can be identified by a combination of morphological (i.e., size, circularity, and clustering) and biochemical (i.e., surface antigens) markers. Here our model for rare breast cancer cells is the MCF7 cell line (breast cancer) spiked in blood. Red blood cells are lysed with a hypotonic lysing agent while MCF7 cells are fixed with formaldehyde and coated with metal beads with a diameter of 1 μm via an antibody to EpCAM (a cell surface molecule that exists on the surface of epithelial cells, but not on the surface of blood cells).

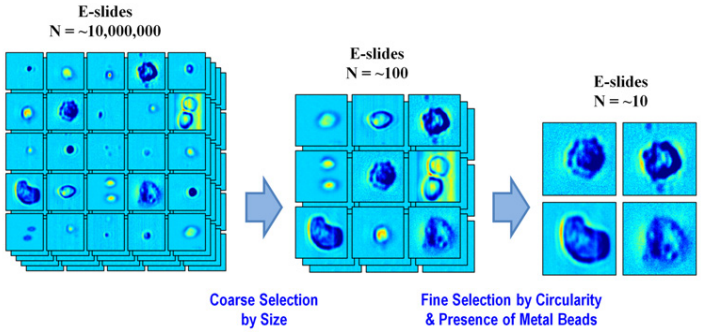


Figure 6: Selection process of the field-programmable digital image processor. The FPGA performs coarse size-based classification and filters out more than 99.9995% of white blood cells while leaving only ~100 false positive events per mL of lysed blood along with true positive events. The CPU then performs fine classification by circularity and presence of metal beads and further down-selects cells by an order of magnitude, leaving true positive events and ~10 false positive events per mL of lysed blood which arise due to image processing artifacts which can further be rejected by human visual inspection.

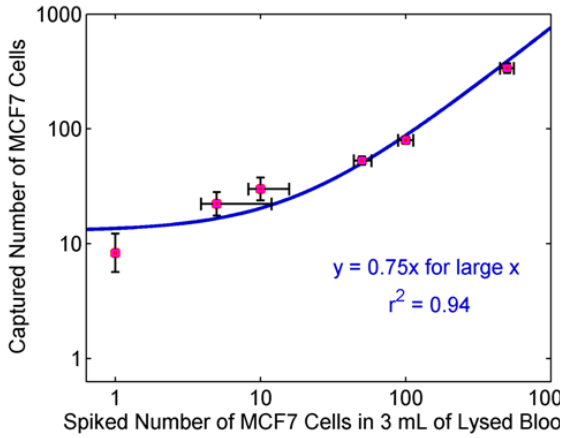


Figure 7: Statistical analysis of the system's capture efficiency for various concentrations. The results indicate that the field-programmable digital image processor can identify extremely rare cells with a high efficiency of 75% (limited by the imperfect coating efficiency and missing smaller MCF7 cells in the FPGA selection process). Here all the measurements were performed with MCF7 cells spiked in buffer containing white blood cells from 3 mL of lysed blood (~80 million white blood cells) at a throughput of 100,000 cells/s. Individual samples were measured four times, establishing that the classification is highly repeatable (indicated by the vertical error bars). Moreover, the correlation of detected MCF7 cells with spiked MCF7 cells is good (r^2 = 0.94). The horizontal deviations can be attributed to several known sources of error in the spiking method including the initial hemacytometer count.

Our observation under a conventional microscope indicates that ~80% of MCF7 cells are coated with 5 – 20 metal beads.

To validate sensitivity for rare cell detection, we performed off-line statistical analysis of white blood cells and coated MCF7 cells. This validation is important as it is required to build a fully trained model for the supervised learning method and hence the algorithms to be implemented on the FPGA and CPU. Figure 5 shows scatter plots of white blood cells and MCF7 cells based on the cell size (i.e., diameter) and presence of surface antigens (i.e., metal beads) using our system. The plots indicate that the system is capable of differentiating most white blood cells and MCF7 cells with high specificity using a single molecular marker. For comparison, Supplementary Figure 3 shows, using a conventional flow cytometer, scatter plots of the same cell types based on forward- and side-scattered light intensities as well as PerCP/Cy5.5-antiCD45 and FITC-antiEpCAM (fluorescence markers) which selectively bind to white blood cells and MCF7 cells, respectively (Appendices). The scattered light intensity plots do not allow for differentiation of the two cell types due to the significant overlap between the two groups. Furthermore, the overlap in the fluorescence intensity limits the specificity of the instrument down to ~0.1% subpopulations – inadequate for identification of circulating tumor cells which are rarer than ~0.1% of the population.

With the fully trained cell classification model, the STEAM flow analyzer can detect MCF7 cells as rare as ~1 in a million in a short period of time. Here we programmed the FPGA and CPU such that the FPGA captures incoming cells and performs size-based cell classification while the CPU performs classification by circularity and presence of metal beads (Supplementary Figure 4). The threshold for the size-based selection is set such that smaller MCF7 cells are also selected at the expense of detecting larger white blood cells (Figure 5). Yet, this process efficiently rejects more than 99.9995% of white blood cells, all residual red blood cells, and all free-floating metal beads, leaving only a small number of false positive events (~100 per mL of lysed blood) along with true positive events (Figure 6). The FPGA's ability to filter out the large number of blood cells and avoid storing the massive amount of digital data enables non-stop real-time operation. The down-selected cell images are stored in the memory (1GB) which can save up to ~4000v images (each of which has 256/v KB, where v is the flow speed in units of m/s). The stored images are then subject to the CPU's selection process in which the cells are further classified by circularity and presence of metal beads (Supplementary Figure 5). This process rejects almost all the false positive events (i.e., multiple, unfocused, or unordered white blood cells) that have survived from the initial selection process on the FPGA while leaving true positive events and a very small number of false positive events (~10 per mL of lysed blood) which arise due to image processing artifacts and can further be rejected by human visual inspection.

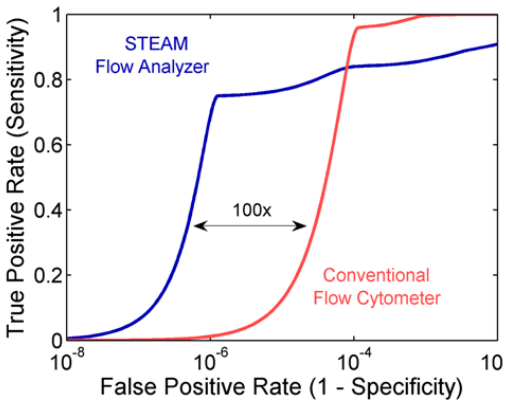


Figure 8: ROC curve analysis of the STEAM flow analyzer in comparison with the conventional flow cytometer. Our method is sufficiently sensitive for detection of ~1 MCF7 cell in a million white blood cells (i.e., a false positive rate of ~10^{-6}) and is two orders of magnitude better in terms of false positive rate than the conventional flow cytometer yet without sacrificing throughput.

Our statistical analysis of the capture efficiency indicates that the field-programmable digital image processor can identify extremely rare cells with a high efficiency of 75% (limited by the imperfect coating efficiency and missing smaller MCF7 cells in the FPGA selection process) (Figure 7). Furthermore, our receiver operating characteristic (ROC) curve analysis of the results indicates that our method is sufficiently sensitive for detection of ~1 MCF7 cell in a million white blood cells and is 100 times better in terms of false positive rate than the conventional flow cytometer (Figure 8), yet without sacrificing throughput. Here all the measurements were performed at a throughput of 100,000 cells/s, corresponding to screening of 10 mL of lysed blood in less than 15 minutes.

Key Research Accomplishments

1. Construction of the optical system including the STEAM camera and its integration with the microfluidic device as well as the optical image amplifier and serializer for high-speed blur-free image acquisition
2. Development of the microfluidic device made of rigid thermoset polyester
3. Development of the real-time optoelectronic image processor that performs high-throughput image-based screening of a large number of cells and identification of rare cells
4. Development of the algorithms implemented on the FPGA and CPU
5. Integration of the hardware and software into a combined system

Reportable Outcomes

1. Successful development of the STEAM flow analyzer
2. Real-time detection of rare MCF7 cells in lysed blood with the analyzer with sensitivity of 75%
3. Demonstration of a false positive rate of 1 in 1,000,000 white blood cells with the analyzer

Conclusion

In summary, we have developed a high-throughput single-cell flow-through image analyzer for real-time image acquisition and screening of a large heterogeneous population of cells with high sensitivity and high statistical accuracy. The throughput is limited by the microfluidic device's tolerance to high pressures caused by high flow rates, not the STEAM camera's image acquisition speed as well as the field-programmable digital image processor's processing speed. We estimate that the throughput can be increased to 200,000 cells/s while retaining reasonable image quality and classification accuracy. Our method can also be combined with a conventional cell sorter to sort out rare cells for further chemical analysis of the cells. While in our proof-of-concept demonstration, we showed high-throughput screening of budding yeast and detection of rare breast cancer cells in blood, our method should also be amenable to other applications in which high-throughput microscopy is required, such as those found in oceanography, energy science, and medicine.

References

[1] K. Goda, K. K. Tsia, and B. Jalali, "Serial time-encoded amplified imaging for real-time observation of fast dynamic phenomena," *Nature*, vol. 458, no. 7242, pp. 1145-1149, Apr. 2009.

[2] D. Di Carlo, "Inertial microfluidics," *Lab on a Chip*, vol. 9, no. 21, pp. 3038-3046, Nov. 2009.

[3] W. Lee, H. Amini, H. A. Stone, and D. Di Carlo, "Dynamic self-assembly and control of microfluidic particle crystals," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 52, pp. 22413-22418, Dec. 2010.

[4] D. Di Carlo, D. Irimia, R. G. Tompkins, and M. Toner, "Continuous inertial focusing, ordering, and separation of particles in microchannels," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 48, pp. 18892-18897, Nov. 2007.

[5] D. R. Gossett and D. Di Carlo, "Particle focusing mechanisms in curving confined flows," *Analytical Chemistry*, vol. 81, no. 20, pp. 8459-8465, Oct. 2009.

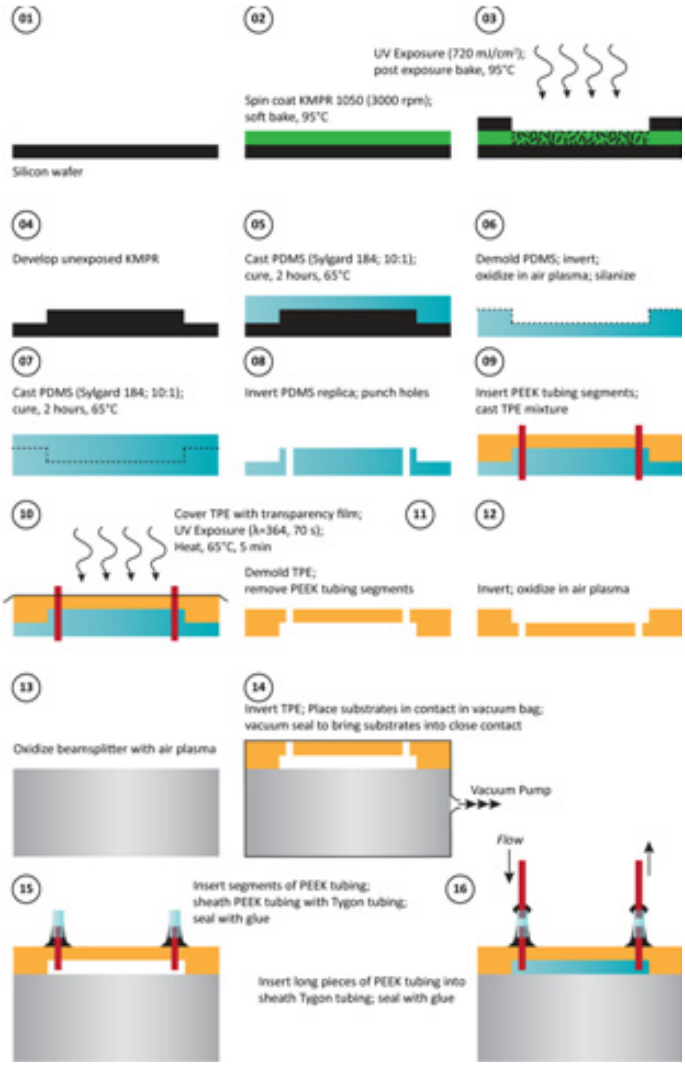
[6] J. S. Kuo et al., "Microfabricating high-aspect-ratio structures in polyurethane-methacrylate (PUMA) disposable microfluidic devices," *Lab on a Chip*, vol. 9, no. 13, pp. 1951-1956, Jul. 2009.

[7] G. S. Fiorini, R. M. Lorenz, J. S. Kuo, and D. T. Chiu, "Rapid prototyping of thermoset polyester microfluidic devices," *Analytical Chemistry*, vol. 76, no. 16, pp. 4697-4704, Aug. 2004.

Appendices

Microfluidic device fabrication

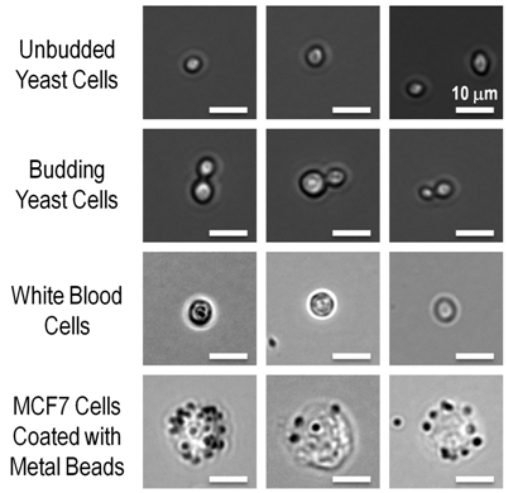
The microfluidic device was fabricated in thermoset polyester (TPE) using standard replica molding methods (Supplementary Figure 1). Our previously reported channels [5] were drawn in AutoCAD (Autodesk, Inc., San Rafael, CA, USA). The design was printed as a photolithography transparency mask at 20,000 dots per inch (dpi). A negative photoresist (KMPR 1050, MicroChem, Corp., Newton, MA, USA) was spun at 3,000 rotations per minute (rpm) on a 4 inch silicon wafer. The wafer was heated on a hot plate at 95°C for 20 minutes then exposed to UV light with a power of 8.0 mW/cm^2 through the transparency mask for 90 seconds. The wafer was baked again at 95°C for 4 minutes then developed in SU-8 developer (MicroChem). The resulting master mold had a feature height of 43 µm. A replica was made by casting Silgard 184 Elastomer Kit (polydimethylsiloxane, PDMS, Dow Corning), mixed according to recipe, over the master.



Supplementary Figure 1: Procedure for the microfluidic device fabrication. Photolithography is used to fabricate a KMP-on-silicon master mold (01-04). A PDMS replica is cured on the master mold and separated from the mold, and a new PDMS mold is then cured on the PDMS replica, resulting in the same polarity as the KMPR-on-silicon master (05-07). Tubing inserts are created and TPE is semi-cured over the PDMS mold (08-10). The TPE replica is removed from the mold, bonded to the substrate, and finally cured (11-14). Tubings are inserted and secured with glue to allow high pressure injections (15-16).

Next, the replica was peeled off the master and PDMS mold of this replica was made as described by Kuo et al [6]. Briefly, the PDMS replica was oxidized in air plasma then silanized with 0.2% octadecyltricholosilane (Sigma-Aldrich) in ethanol for 15 minutes. It was then rinsed with ethanol and heated to 65°C for 30 minutes to prevent swelling. PDMS was then cast on the PDMS replica to create a mold with the same polarity as the master. It was degassed and cured at 65°C for 2 hours. This mold was separated from the replica and sonicated in isopropanol for 5 minutes, sonicated in deionized H_2O for 5 minutes, and baked at 65°C for at least 30 minutes. Holes were punched for tubing connectors.

TPE channels were fabricated following the protocol of Fiorini et al [7]. 0.10 g of the UV-photoinitiator (2,2-dimethoxyphenylacetophenone, Irgacure 651, Sigma-Aldrich) was dissolved in 0.25 g of styrene monomer (Sigma-Aldrich). 10 g of resin (Polylite 32030-10, Reichhold, Inc.) and 0.09 g of methyl-ethyl-ketone-peroxide catalyst (MEKP, Crompton Corp.) were added. The mixture was stirred then degassed for 5 minutes. PEEK (Polyetheretherketone) tubing segments were inserted into the PDMS mold to provide a molded region for tubing connectors to be inserted later. The TPE mixture was cast over the PDMS mold to a thickness of 3-4 mm and covered with transparency film (3M), cut to an appropriate size to ensure a flat, top surface during curing and avoid inhibition of cross-linking reagents by air. The TPE was exposed to UV for 70 seconds (λ = 364 nm, 400 W, Dymax Model 2000 Flood) and immediately heated at 65°C for 5 minutes. The PEEK tubing segments were removed from the chip and the TPE was carefully pulled away from the mold while still slightly pliable. Both the cured TPE and a 2" BK7 beamsplitter were oxidized in air plasma (1 minute at 200 mTorr) and brought into close contact with a vaccum sealer.



Supplementary Figure 2: CCD images of stationary cells. Images of cells of various species were taken by a research-grade CCD camera. The exposure time (shutter speed) and pixel number of the CCD camera are 17 ms and 1280 x 1024, respectively.

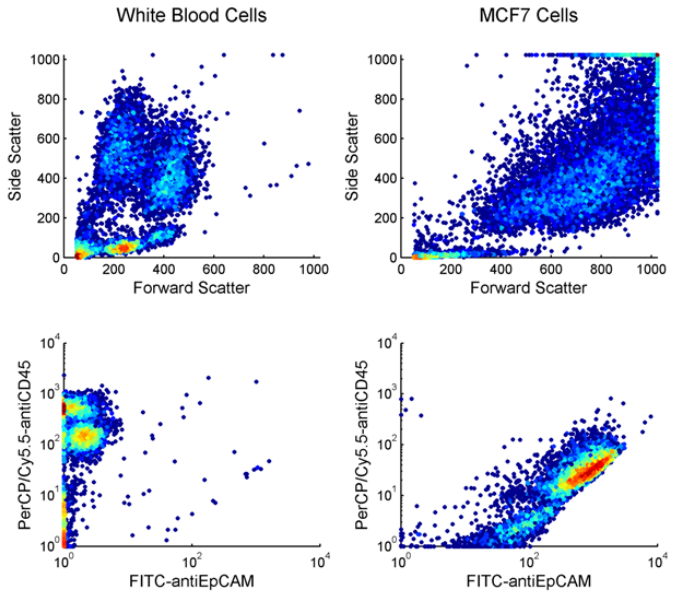
The TPE-beamsplitter channels were left to cure and harden at room temperature for 1 day. Clean segments of PEEK tubing were inserted into the molded tubing connection holes and sheathed with Tygon tubing segments. The connections were sealed at room temperature with acrylate glue (Krazy Glue®) and UV curing NOA81 (ThorLabs). Longer segments of PEEK tubing were inserted into the free end of the Tygon tubing segments. A summary of the microfluidic device's specifications is shown in Supplementary Table 2.

Illumination and imaging optics of the STEAM camera

The optical source is a mode-locked femtosecond pulse fiber laser with a pulse repetition rate of 36.7 MHz. After supercontinuum generation in a highly nonlinear fiber and band-pass filtering, a nearly flat spectral shape with ~20 nm bandwidth centered at 1590 nm is produced for target illumination. A pair of diffraction gratings with 1100 lines/mm spatially disperses the pulses along a 1D line, producing 1D rainbow pulses.

Amplified dispersive Fourier transformer

The amplified dispersive Fourier transformer (ADFT) consists of a dispersive fiber with -1200 ps/nm dispersion pumped by four continuous-wave lasers with ~100 mW at 1470 nm, 1480 nm, 1480 nm, and 1490 nm for distributed Raman amplification. Before the ADFT, an erbium-doped fiber amplifier (EDFA) is used as a pre-amplifier. The EDFA and ADFT produce a net gain of up to 13 dB and 17 dB (depending on signal-to-noise ratio), respectively. The specifications of the STEAM camera can be estimated from the parameters of its components. First, the number of 1D image pixels on the target (N) is found from the total dispersion in the dispersive fiber (D = -1200 ps/nm), the optical bandwidth ($\Delta\lambda$ = 20 nm), and the sampling rate of the ADC in the field-programmable digital image processor (f_{dig} = 7 GS/s) to be $N = D \cdot \Delta\lambda \cdot f_{dig}$ = 168 (which is reduced down to 128 by removing the frame edges) whereas the number of 1D resolvable points is about 200 from the spectral resolution of the ADFT process. Second, the spatial resolution of the STEAM camera is found to be about 1.4 μm from the diffraction gratings and imaging optics which consist of a spherical lens with 850 mm focal length and an objective lens with a numerical aperture of 0.4. Finally, the shutter speed (exposure time) is estimated from the bandwidth of each rainbow pulse (20 nm / ~200) and the time-bandwidth product to be 27 ps (assuming that each subpulse of the rainbow pulse is near transform limited).



Supplementary Figure 3: ***Conventional flow cytometer's scatter plots.*** *Shown are events of white blood cells and MCF7 cells based on forward- and side-scattered light intensities as well as PerCP/Cy5.5-antiCD45 and FITC-antiEpCAM (fluorescence markers) which selectively bind to white blood cells and MCF7 cells, respectively (See Methods). The total number of events is ~10,000 for both cell types. The scattered light intensity plots are clearly inadequate for differentiation of the two cell types due to the significant overlap between the two groups. Furthermore, overlaps in the fluorescence scatters due to unwanted autofluorescence and nonlinear optical processes limit the specificity of the instrument down to ~0.1% subpopulations – inadequate for identification of a cell type which is rarer than ~0.1% of the population.*

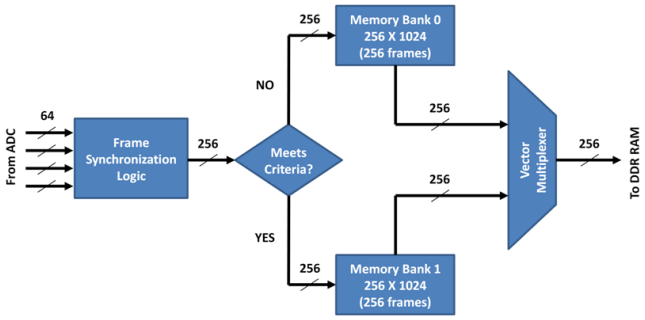
CMOS and CCD cameras

The CMOS camera used in Figure 4 is Phantom v7.3 (Vision Research). It has an exposure time (shutter speed) of 1 µs and a pixel number of 32 x 32. The CCD camera used for Supplementary Figure 2 is DS-Qi1 with a DS-U2 PC-based control unit (Nikon). It has an exposure time (shutter speed) of 17 ms and a pixel number of 1280 x 1024.

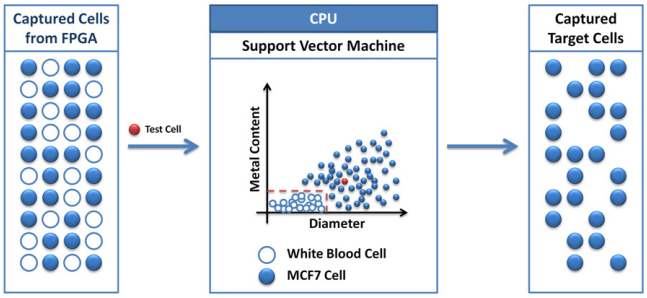
Architecture of the field-programmable digital image processor

The field-programmable digital image processor employs a programmable multi-stage decision-making model that can be tailored, depending on the application. It consists of the ADC, FPGA, memory circuit, and CPU. The structure and function of the FPGA (coarse) and CPU (fine) classification processes are schematically shown in Supplementary Figures 4 and 5.

The FPGA classification process is described as follows. First, the digitized output of the ADC sampled at 7 GS/s by time interleaving is inputted to the on-board Xilinx Virtex-6 FPGA device (LX240T). The FPGA runs at a frequency of 200 MHz and receives 32 samples of time-interleaved data at every clock cycle. The incoming data is synchronized in the FPGA by selecting four clock cycles of data from the first positive sample point value in the rising edge of the pulse (i.e., the 1D image frame) which forms one frame of the image (i.e., 128 sample points per frame). Initially, the fourth incoming frame is selected as the reference frame by the FPGA and it is done while PBS is passed through the microfluidic channel. Once the reference frame has been selected by the FPGA, the incoming frames are checked for meeting criteria of the target cell. The criteria check consists of comparing the difference between the value of sample points in the incoming frame and the corresponding points in the reference frame to a threshold value which matches to the value observed at the center of a typical coated MCF7 cell, which is different from the typical white blood cell. 2048/v frames divided by the flow speed (where v is the flow speed in units of m/s) are used to represent a MCF7 cell. Therefore, once the criteria are met, 1024/v frames stored until the sample point where the criteria are satisfied and the next incoming 1024/v frames constitute the complete MCF7 cell image to be stored into the on-board 1 GB DDR RAM.



Supplementary Figure 4: Block diagram of the hardware implemented on the FPGA for the rare cell detection. The FPGA performs size-based cell classification. This process filters out nearly 99.9995% of white blood cells, all residual red blood cells, and all free-floating metal beads, leaving only ~100 false positive events along with true positive events.



Supplementary Figure 5: SVM implemented on the CPU for the rare cell detection. This cell filtration process is performed off-line on the captured cells by the FPGA to differentiate similar looking MCF7 cells and white blood cells based on more delicate differences of their images. The cell classification program uses a SVM to classify the images. The optimized SVM utilizing size, metal content, and circularity (including clustering) can identify rare MCF7 cells with a false positive rate of ~1 in a million.

While criteria checking on an incoming quarter frame (32 samples) is taking place at a clock cycle, the quarter frames are stored into the memory bank 0 as well which is a 256 x 4096/v block RAM in the FPGA which can store 1024/v frames as 4096/v quarter frames which are 256 bits wide. The incoming frame from the point where the MCF7 cell criteria are met gets stored into the memory bank 1 of 256 x 4096/v block RAM in the FPGA while the cell image frames stored in the bank 0 is transferred to the on-board 1 GB DDR RAM. After completion of the frame transfer from bank 0, the input to the external DDR RAM is multiplexed to the output of memory bank 1 to transfer the frames stored in it and simultaneously the incoming frames are checked for the MCF7 cell and gets written to memory bank 0 as explained before.

After the initial filtration process on the FPGA captures nearly all of the MCF7 cells along with occasional suspicious looking white blood cells, their images are transferred to the CPU for further processing. A cell classification algorithm is prepared by studying the images of a large number of cells from each cell type to identify key differences in their signatures. This further cell filtration is implemented off-line on the captured cells to differentiate similar looking MCF7 cells and white blood cells based on more delicate differences of their images. This very efficient, extremely accurate image processing algorithm would require more processing time and needs to be implemented as a post-processing step on a more powerful processor than the FPGA. It takes advantage of a large library of cells each marked as belonging to one of two categories: white blood cells and MCF7 cells. The supervised learning algorithm analyzes the training data and produces a model that maps each image into a two-dimensional space such that cells from different categories are separated in this space. New examples from mixed blood samples that are captured and pre-filtered by the FPGA are then mapped into that same space and predicted to belong to a category based on their location in this space.

Parameter	Value
Throughput (Corresponding Flow Speed)	50,000 particles/s (2 m/s) 100,000 particles/s (4 m/s) 200,000 particles/s (8 m/s)
Frame Rate	36.7 MHz (Line Scan Mode)
Shutter Speed	27 ps
Flow Technology	Inertial Focusing and Ordering
Microfluidic Device	Thermoset Polyester
Flow Speed	Up to 8 m/s
Optical Image Gain	Up to 30 dB
Sampling Rate	7 GS/s
Pixel Number (Corresponding Flow Speed)	• 128 x 1024 (2 m/s) • 128 x 512 (4 m/s) • 128 x 256 (8 m/s)
Field of View	30 x 30 μm^2
Laser Center Wavelength	1590 nm
Optical Bandwidth	20 nm
Illumination Power	500 μW
Resolution	1.4 μm
Image Analysis	Online Real-Time

Supplementary Table 1: Specifications of the STEAM flow analyzer. The throughput is limited by the microfluidic device's tolerance to high pressures caused by high flow rates, not the STEAM camera's shutter speed and frame as well as the real-time optoelectronic image processor's processing speed. For reasonable image quality and screening accuracy, the system can operate at a throughput of 200,000 cells/s.

The cell classification program uses a SVM to classify the images. During the training process, the SVM algorithm defines and optimizes key features based on size, metal content, and circularity of cells. Metal content is correlated to the introduction of regions with strong shadows on the cell image due to attachment of the metal beads. It is a good measure of surface antigens, which are specific to MCF7 cells. On the other hand, the exact size of the cell from its image serves as an extra measure of differentiating the two cell types, as MCF7 cells are on average larger than white blood cells. Furthermore, circularity of cells is also evaluated as another measure of cell differentiation since white blood cells are typically circular or symmetrically shaped while MCF7 cells are irregularly or asymmetrically shaped or clustered with a few to several MCF7 cells. The optimized SVM utilizing size, metal content, and circularity can identify rare MCF7 cells with a false positive rate of ~1 in a million.

Parameter	Value
Width of Microfluidic Channel at Interrogation Point	112 μm
Height of Microfluidic Channel at Interrogation Point	43 μm
Length of Microfluidic Channel	2.2 cm
Volumetric Flow Rate (Corresponding Mean Channel Velocity)	0.6 mL/min (2 m/s) 1.2 mL/min (4 m/s) 4.8 mL/min (8 m/s)
Channel Material	Thermoset Polyester
Maximum Pressure	≥150 psi (Limit of Pressure Sensor)
Young's Modulus	1.2 GPa
Light Transmittance	~90%
Swelling Ratio in Ethanol	1.0 (No Swelling)

Supplementary Table 2: Specifications of the microfluidic device. To ensure stability in the real-time cell classification, the microchannel was fabricated by replica molding in thermoset polyester due to its ease of fabrication, high stiffness, ability to sustain high pressures, and resistance to ethanol which helps channel washing. Among these advantages, channel rigidity is especially critical to avoid undesirable fluid dynamic and optical effects which result if the channel deforms.

Flow cytometry experiment

MCF7 cells were harvested and fixed in 0.5% formaldehyde for 10 minutes, washed and resuspended in PBS. Whole blood was mixed with red blood cell lysis buffer (Roche) at a ratio of 2 (lysis buffer) to 1 (whole blood) and incubated for 10 minutes. The remaining white blood cells were resuspended once in lysis buffer, then in PBS. Both MCF7 cell and white blood cell suspensions were incubated with FITC-antiEpCAM and PerCP/Cy5.5-antiCD45. Unbound dyes were removed after 30 minutes and cells were washed and resuspended in PBS. The pure MCF7 and white blood cell suspensions and two mixtures were analyzed with the FACSCalibur system in the Flow Cytometry Core Laboratory at UCLA.

Rare cell detection experiment

The initial concentration of cells which were used for the rare cell detection measurements was determined by a hemacytometer. Four counts were performed and the mean and standard deviation were calculated. Clusters of cells were counted as a single cell. The fate of these clusters in the microfluidic device is unclear; clusters could be filtered out upstream, traverse the channel as a cluster, or break up into smaller clusters or individual cells. The coated MCF7 cells were diluted into 3 mL of lysed blood. The resulting concentrations were 0, 5, 10, 50, 100, and 500 coated MCF7 cells per 3 mL. The standard deviation of the initial count is expected to decrease by the dilution factor for each spiked sample. Other sources of error, such as in the process of pipetting, are more difficult to quantify due to the lack of comparable technology for detecting extremely rare events or low concentrations of cells. The fit function used in Figure 8 is y = ax + b, where x and y are the spiked and captured numbers of MCF7 cells, respectively, a is the capture efficiency, and b is the contamination. For large x, the fit function asymptotically approaches y = ax.

Parameter	Value
FPGA Device	Xilinx Virtex-6 LX240T
FPGA Clock Frequency	200 MHz
Frame Size	128 bytes (Pixels/Samples)
Samples Received in One Clock Cycle	32 bytes (Quarter Frame)
E-slide Size Captured in FPGA	2048/v Frames Consisting of 1024/v Frames in Two Memory Banks of Block RAM (v = Flow Speed in Units of m/s)
External Memory (DDR RAM) for Storing E-slides	1 GB
Memory Size of Single E-slide	256/v KB
Maximum Number of E-slides Stored in DDR RAM	4,000v

Supplementary Table 3: Specifications of the field-programmable digital image processor. The FPGA device used in our system is a state-of-the-art FPGA (Xilinx Virtex-6 LX240T). The E-slide size captured in the FPGA can be tailored, depending on the size of the particles of interest.